

Leveraging Frequency Analysis for Deep Fake Image Recognition

Year: 2020 | Citations: 706 | Authors: J. Frank, Thorsten Eisenhofer, Lea Schönherr, Asja Fischer, D. Kolossa

Abstract

Deep neural networks can generate images that are astonishingly realistic, so much so that it is often hard for humans to distinguish them from actual photos. These achievements have been largely made possible by Generative Adversarial Networks (GANs). While deep fake images have been thoroughly investigated in the image domain - a classical approach from the area of image forensics - an analysis in the frequency domain has been missing so far. In this paper, we address this shortcoming and our results reveal that in frequency space, GAN-generated images exhibit severe artifacts that can be easily identified. We perform a comprehensive analysis, showing that these artifacts are consistent across different neural network architectures, data sets, and resolutions. In a further investigation, we demonstrate that these artifacts are caused by upsampling operations found in all current GAN architectures, indicating a structural and fundamental problem in the way images are generated via GANs. Based on this analysis, we demonstrate how the frequency representation can be used to identify deep fake images in an automated way, surpassing state-of-the-art methods.